

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2002, 21:34:49 ; Search time 1445.89 Seconds
(without alignments)
15540.019 Million cell updates/sec

Title: US-09-635-521A-1
Perfect score: 1362
Sequence: 1 atggcttaccacagctcc.....ttcagagcatgaatttga 1362

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
To: number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	100.0	1362	9 AF034633	AF034633 Homo sapi
2	857	62.9	174494	2 AC068744	AC068744 Homo sapi
3	508.6	37.3	129676	9 AC079773	AC079773 Homo sapi
4	497.6	36.5	1890	6 AX136281	AX136281 Sequence
5	285	20.9	1797	10 AB041649	AB041649 Mus muscu
6	236.6	17.4	587	6 AX136698	AX136698 Sequence
7	165.2	12.1	1254	22 E11480	E11480 cdna encod
8	164.8	12.1	4131	9 HSNEDURA	X70070 H.sapiens m
9	158.6	11.6	444	6 AX150120	AX150120 Sequence
10	150.4	11.0	110000	2 AL357872	AL357872 Homo sapi
11	150.4	11.0	114642	6 AL357033	AL357033 Human DNA
12	133.8	9.8	3917	10 AB017027	AB017027 Mus muscu
13	124.8	9.2	1239	9 AK026195	AK026195 Homo sapi
14	123.2	9.0	170695	2 AL391318	AL391318 Homo sapi
15	123.2	9.0	189476	2 AL157394	AL157394 Homo sapi
16	97.6	7.2	1569	9 HSNTR2REC	Y10148 H.sapiens m
17	96.2	7.1	1676	5 AF082210	AF082210 Spheroicoid
18	95.8	7.0	1342	6 E29334	E29334 Novel human
19	89.4	6.6	729	6 AX139107	AX139107 Sequence
20	89.4	6.6	801	6 AX147772	AX147772 Sequence
21	89.4	6.6	1239	9 AF242874	AF242874 Homo sapi
22	89.4	6.6	1248	9 AB041228	AB041228 Homo sapi
23	89.4	6.6	1248	9 AF292402	AF292402 Homo sapi
24	89.4	6.6	1298	9 AF272363	AF272363 Homo sapi
25	89.4	6.6	1594	6 AX109242	AX109242 Sequence
26	89.4	6.6	1594	6 AX109244	AX109244 Sequence
27	89.4	6.6	1658	6 AX109234	AX109234 Sequence
28	89.4	6.6	1658	6 AX109236	AX109236 Sequence
29	89.4	6.6	1658	6 AX109238	AX109238 Sequence
30	89.4	6.6	1658	6 AX109240	AX109240 Sequence
31	89.4	6.6	215088	2 AC008571	AC008571 Homo sapi
32	86.4	6.3	1554	10 MMU51908	U51908 Mus muscu
33	85.4	6.3	1529	10 RNMTR2REC	X87121 R.norvegicu
34	83.6	6.1	6787	9 AF369786	AF369786 Homo sapi
35	83.6	6.1	145593	9 AC069523	AC069523 Homo sapi
36	82	6.0	870	9 HS060181	U60181 Human growt
37	82	6.0	1088	6 AR156353	AR156353 Sequence
38	82	6.0	1101	6 AX154584	AX154584 Sequence
39	82	6.0	1101	6 HS060179	U60179 Human growt
40	82	6.0	1122	6 AR156354	AR156354 Sequence
41	80.4	5.9	145272	2 AC016938	AC016938 Homo sapi
42	78.6	5.8	836	6 AR156355	AR156355 Sequence
43	77.2	5.7	1092	6 AR156357	AR156357 Sequence
44	77.2	5.7	1350	6 AB001982	AB001982 Rattus no
45	77.2	5.7	3129	6 AR156356	AR156356 Sequence

ALIGNMENTS

RESULT 1
AF034633 1362 bp mRNA PRI 13-JUL-1998
LOCUS Homo sapiens orphan G protein-coupled receptor (GPR39) mRNA,
DEFINITION complete cds.

ACCESSION AF034633
VERSION AF034633.1 GI:2654160

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

1 (bases 1 to 1362)
McKee,K.K., Tan,C.P., Palyha,O.C., Liu,J., Feighofer,S.D.,
Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,H.
Cloning and characterization of two human G protein-coupled
receptor genes (GPR38 and GPR39) related to the growth hormone
secretagogue and neurotensin receptors
Genomics 46 (3), 426-434 (1997)

SUMMARIES

REFERENCE 2 (bases 1 to 1362)
 McKee, K.K., Tan, C.P., Palyha, O.C., Liu, J., Feighner, S.D.,
 Hreniuk, D.L., Smith, R.G., Van Der Ploeg, L.H.T. and Howard, A.D.
 TITLE Direct Submission
 JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co.,
 Inc., PO Box 2000, Rahway, NJ 07065, USA
 FEATURES
 Location/Qualifiers
 1..1362
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q21-22"
 /dev_stage="fetal"
 /tissue_type="brain"
 1..1362
 /gene="GPR39"
 1..1362
 /gene="GPR39"

hormone secretagogue and neurotensin receptors"
 /product="GPR39"
 /protein_id="AAC26082.1"
 /db_xref="GI:2654161"
 /translation="MASPSPGSDSQIIDSHVPEPEVAWIKILIVLYLIFVWG
 LIGNSATIRVYOLOKGYLOKEVTDHMSLACSDILVPLIGMMEYSIIMPLPTS
 SYTLSCILHTEIFACSYATLLHYLTLSFERYTACIPREYKASGCOYKILIGYV
 VTSLALPLIFAMGTREYPLVNPSSHGLICNSTRHHQPERSNSTICNLSRMT
 VFOSSIRGARVYLVILVLSVAFKCMNMVOMLSOKSLGSTRPOLRSESEEST
 ARROTIIIFELIIVTLAVCMPPNDRIRIMAAKPKHDWTSYRAYVILILPSEETFEY
 LSVYINPLTVSSQDFRVEVOYLCCRLSIQHANHEKRLRVAHSHSTDSARVORPL
 LFRSRROSSARTEKIFLSTFOSEAEPOSSQSLSELEPNAGKAPNANBAENGFOE
 HEV"

BASE COUNT 263 a 435 c 362 g 302 t
 ORIGIN

Query Match 100.0%; Score 1362; DB 9; Length 1362;
 Best Local Similarity 100.0%; Pred. No. 4.8e-262;
 Matches 1362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atggcttaccacacccctccgggagcagtcgctcccaatcatgatacagtcgtc 60
 1 atggcttaccacacccctccgggagcagtcgctcccaatcatgatacagtcgtc 60
 61 ccgagattgagtgccacacacacacacacacacacacacacacacacacacac 120
 61 ccgagattgagtgccacacacacacacacacacacacacacacacacacacac 120
 121 ttgctgagtgccac 180
 121 ttgctgagtgccac 180
 181 aaagagatctgaggaagagagagagagagagagagagagagagagagagagag 240
 181 aaagagatctgaggaagagagagagagagagagagagagagagagagagagag 240
 241 ttgctgagtgccac 300
 241 ttgctgagtgccac 300
 301 acgtcagatcagac 360
 301 acgtcagatcagac 360
 361 gctacagtgctgac 420
 361 gctacagtgctgac 420
 421 ttcagagacagagctgctgagagacacacacacacacacacacacacacacac 480
 421 ttcagagacagagctgctgagagacacacacacacacacacacacacacacac 480

481 gtacactccgcccgtgagacacacacacacacacacacacacacacacacacac 540
 481 gtacactccgcccgtgagacacacacacacacacacacacacacacacacacac 540
 541 gtgagctgcccac 600
 541 gtgagctgcccac 600
 601 cagccagagac 660
 601 cagccagagac 660
 661 cagccagagac 720
 661 cagccagagac 720
 721 atgctgagac 780
 721 atgctgagac 780
 781 acgagcctccgagac 840
 781 acgagcctccgagac 840
 841 accatcactccgagac 900
 841 accatcactccgagac 900
 901 attcgagagac 960
 901 attcgagagac 960
 961 gctgagac 1020
 961 gctgagac 1020
 1021 ccgctcctgagac 1080
 1021 ccgctcctgagac 1080
 1081 tgcgcctgctgagac 1140
 1081 tgcgcctgctgagac 1140
 1141 accacagagac 1200
 1141 accacagagac 1200
 1201 tctgcaagagac 1260
 1201 tctgcaagagac 1260
 1261 tctgcaagagac 1320
 1261 tctgcaagagac 1320
 1321 aattctgctgagac 1362
 1321 aattctgctgagac 1362

RESULT 2
 AC068744 174494 bp DNA HTG 04-OCT-2000
 LOCUS
 DEFINITION Homo sapiens chromosome 2 clone RP11-666P13 map 2, WORKING DRAFT
 AC068744
 AC068744.2 GI:10567984
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 17494)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 2, clone RP11-666P13
Unpublished
2 (bases 1 to 17494)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Bouknight,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Deatello,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hages,B., Heatford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R., Meldrum,J., Menes,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J., Murphy,I., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Submitted (08-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA.
On Oct 4, 2000 this sequence version replaced g1:7717151.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Center code: MTRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L9278
Center clone name: 666_P_13

----- Summary Statistics -----
Sequencing vector: M13: M7815: 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 163890 bases at least Q40
Consensus quality: 169058 bases at least Q30
Consensus quality: 171182 bases at least Q20
Insert size: 170000: agarose-fp
Insert size: 172394: sum-of-ctnigs
Quality coverage: 4.9 in Q20 bases; sum-of-ctnigs
Quality coverage: 4.8 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 17735: contig of 17735 bp in length
* 17736 17835: gap of 100 bp
* 17836 45012: contig of 27177 bp in length
* 45013 45112: gap of 100 bp
* 45113 46230: contig of 1118 bp in length
* 46231 46330: gap of 100 bp
* 46331 47605: contig of 1275 bp in length
* 47606 47705: gap of 100 bp
* 47706 49342: contig of 1637 bp in length
* 49343 49442: gap of 100 bp
* 49443 51470: contig of 2028 bp in length
* 51471 51570: gap of 100 bp

FEATURES

source

51571 53723: contig of 2153 bp in length
* 53724 53823: gap of 100 bp
* 53824 56799: contig of 2976 bp in length
* 56800 56899: gap of 100 bp
* 56900 60103: contig of 3204 bp in length
* 60104 60203: gap of 100 bp
* 60204 64890: contig of 4687 bp in length
* 64891 64990: gap of 100 bp
* 64991 70638: contig of 5648 bp in length
* 70639 70738: gap of 100 bp
* 70739 76977: contig of 6239 bp in length
* 76978 77077: gap of 100 bp
* 77078 82744: contig of 5667 bp in length
* 82745 82844: gap of 100 bp
* 82845 90681: contig of 7837 bp in length
* 90682 90781: gap of 100 bp
* 90782 98252: contig of 7471 bp in length
* 98253 98352: gap of 100 bp
* 98353 107826: contig of 9474 bp in length
* 107827 107926: gap of 100 bp
* 107927 117679: contig of 9753 bp in length
* 117680 117778: gap of 100 bp
* 117780 129303: contig of 11524 bp in length
* 129304 129403: gap of 100 bp
* 129404 140999: contig of 11596 bp in length
* 141000 141099: gap of 100 bp
* 141100 151374: contig of 10275 bp in length
* 151375 151474: gap of 100 bp
* 151475 174223: contig of 22749 bp in length
* 174224 174323: gap of 100 bp
* 174324 174494: contig of 171 bp in length.
Location/Qualifiers
1. 174494
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone.lib="RP11-666P13"
/clone.lib="RP11-11 Human Male BAC"
1. 17735
/note="assembly-fragment"
clone_end:SP6
vector_side:left
17836..45012
/note="assembly-fragment"
45113..46230
/note="assembly-fragment"
46331..47605
/note="assembly-fragment"
47706..49342
/note="assembly-fragment"
49443..51470
/note="assembly-fragment"
51571..53723
/note="assembly-fragment"
53824..56799
/note="assembly-fragment"
56900..60103
/note="assembly-fragment"
60204..64890
/note="assembly-fragment"
64991..70638
/note="assembly-fragment"
70739..76977
/note="assembly-fragment"
77078..82744
/note="assembly-fragment"
82845..90681
/note="assembly-fragment"
90782..98252
/note="assembly-fragment"
98353..107826
/note="assembly-fragment"

```

misc_feature      107927..117679
                  /note="assembly_fragment"
misc_feature      117780..129303
                  /note="assembly_fragment"
misc_feature      129404..140999
                  /note="assembly_fragment"
misc_feature      141100..151374
                  /note="assembly_fragment"
misc_feature      151475..174223
                  /note="assembly_fragment"
misc_feature      174324..174494
                  /note="assembly_fragment"

```

```

BASE COUNT      50891 a 36548 c 36228 g 48718 t 2109 others
ORIGIN

```

```

Query Match      62.9%; Score 857; DB 2; Length 174494;
                  Local Similarity 100.0%; Pred. No. 1.2e-161;
                  Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 atggcttaccagcgtccggcgagtgatgctcccaatcatgacagtcgtc 60
    |||
Db 28320 ATGGCTTACCAGCCTCCGGCGAGTGTCTCCAAATCATGATCAGTCATGTC 28379

QY 61 ccgagttgagtgccacgtgatacaaatcaacctattctgtgtacatgatac 120
    |||
Db 28380 CCCGAGTTGAGTGGCCACCGATCAAAATCAACCTTATTCTGTCTACTGATATC 28439

QY 121 ttctgtgagtggtctctgtgggaacagcgccacattgggttcaaccaagtgctgagaag 180
    |||
Db 28440 TTCTGTGAGTGGGCTTCTGGGAAACGCCCAACATTCGGGTACCCAGTCTGCAGAG 28499

QY 181 aaagatacttcgaaagagtgacagacacatggtgagtttggtctgtcggaac 240
    |||
Db 28500 AAAGATATCTTCGAAAGAGAGGTGACACACATGTAAGTGTGGCTGTGAGATC 28559

QY 241 ttgtgttctctcatcgatgccatgaggttctacagcatcattgaaatccctgac 300
    |||
Db 28560 TTGTGTCTCTCATCGATGCCATGAGTGTCAACATCATCTGGAATCCCTGATCC 28619

QY 301 acgtcaacttacaacctgtctctgcaagctgacacattctcttcgagagcctgacgtac 360
    |||
Db 28620 ACGTCACTACACCTCTCTCTGCAAGCTGCACACTTCTCTTGAAGCCTGCAGACTAC 28679

QY 361 gctacgctgcgcagctgtctacactacgtttgagcgtcactgcgcacatgtgcacccc 420
    |||
Db 3680 GCTACGCTGCAGCTGTCTACACTGACTTTGAGGCTTACATGCTGTCACCCC 28739

QY 421 ttcaagtaaacgctgtgtcgggaaccttgcaagtgaaagtctgattgtctgtctg 480
    |||
Db 28740 TTCAAGTAAAGCGTGTGTGCGGACCTTGCAAGTGAAGTGTGATGGTGTCTGG 28799

QY 481 gtcaacttccggcgtgtgagccttgccttgtgttgcacatggttactggttaacctctg 540
    |||
Db 28800 GTCACTCTCCGCTGTGTGTGCACTGCTGCTGTGTTGCAATGAGTACTGATGATCCCTG 28859

QY 541 gttgaagctgcagcagcagcaggtgtctactgtcaacgcttccacacccgcaacagag 600
    |||
Db 28860 GTGAAGCGGCCACGACCGGGGCTCACTTGCACACGCTCCAGACCCGCGCACAGAG 28919

QY 601 cagcccgagacctcaaatatgtcatctgttacaaccttccacgcgctggaacgtgttc 660
    |||
Db 28920 CAGCCCGAGACCTCAATATGTTCATCTGTACCAACTTCCAGACCCGCTGAGACGTTTC 28979

QY 661 cagtcacgacatcttcgagcgtctgtgttactactggtgtctgtctctcgtagccttc 720
    |||
Db 28980 CAGTCACGACATCTTCGCGCTGTGTGTACTCTGTGTCTGCTGCTCTCCGTAAGCCTTC 29039

QY 721 atgtgtcgggaacatgacgtgtgtctatgaaagcagaagaggtcgtgcgcgagagc 780
    |||
Db 29040 ATGTGTCTGGGAACATGATGAGTGTCTATGAAAGCCAGAAAGGCTCGCTGGCGGGGGC 29099

```

```

QY 781 acggggtctccgacgtgaggaagtcgcagagagcagagccagagagcag 840
    |||
Db 29100 ACGGGGCTCCGCGAGTGTGAGAACTCGAGAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 29159

QY 841 accatcatcttcctgag 857
    |||
Db 29160 ACCATCATCTTCTCCGAG 29176

```

```

RESULT 3
AC079773      129676 bp      DNA      PRI      09-AUG-2001
LOCUS      Homo sapiens chromosome 2 clone RP11-258B17, complete sequence.
DEFINITION      AC079773
VERSION      AC079773.8 GI:15145561
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS      Waterston,R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (10-SEP-2000) Genome Sequencing Center, Washington
                University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                MO 63108, USA

```

```

REFERENCE
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (09-AUG-2001) Genome Sequencing Center, Washington
                University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                MO 63108, USA
COMMENT
On Aug 9, 2001 this sequence version replaced gi:14488388.

```

```

FEATURES
source
    1..129676
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="2"
    /clone="RP11-258B17"
    ----- Project Information -----
    Center project name: H.NH0258817

```

```

BASE COUNT      35366 a 27626 c 29232 g 37452 t
ORIGIN

```

```

Query Match      37.3%; Score 508.6; DB 9; Length 129676;
Best Local Similarity 99.2%; Pred. No. 4.3e-92;
Matches 511; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 848 tcttcttgagctgtatgttgtgacattggtcgatgtgatagtcccaacagattcgga 907
    |||
Db 119552 TCGGCCAGGGCGATGTGTGACATTGGCGCTATGCTGATGCCAAGACGATTCGGA 119611

QY 908 ggaatcagctgcgagcaaaccaacagacgactgagaggtctactctcggtcgataga 967
    |||
Db 119612 GGAATCATGTGCTCGGCGCAAAACCAAGACGACTGGAACAGAGTCTACTTCCGGCGGTACA 119671

QY 968 tgatctctctcccttctctgagagagtttttctactcaagctcgtgtatcaacccgtctcc 1027
    |||
Db 119672 TGATCTCTCTCCCTTCTCGGAGACGTTTTTCTACTCAGCTCGGTCTATCAACCGCTCC 119731

```

Qy	1028	tgtaacagtgatccctcgcaacagattcgggggagtggttcctcgtcaagagtgctctgtgcgcgc	1087
Db	119732	tgtaacaggtgtctctccacagcagttttggcgggggtgttcctgcacagctgtcgtgcgcgc	119791
Qy	1088	tgtcgtctcagcagcccaacaacacagagaagcgctcgtcgtctacatgycgaatccacaacg	1147
Db	119792	tgatcgctgacagcagcccaacacagagaaaggcgtgcgcctgacatgccaactccacacgc	119851
Qy	1148	acagcgccgcgttctgtgcagcgcccgctgtcctcttcgcgtcccgcgcgcaagtcctctgcaa	1207
Db	119852	acagcgccgcgttctgtgcagcgcccgctgtcctcttcgcgtcccgcgcgcaagtcctctgcaa	119911
Qy	1208	ggaggaactgtaggaatttctttaagcactttcaagcgagcgcccgagccccagtcataagt	1267
Db	119912	ggaggaactgtaggaatttctttaagcactttcaagcgagcgcccgagccccagtcataagt	119971
Qy	1268	cccgatcatctgaagtcctcgaagtcactaagcccaactcagcggcgaaaccgcgaattctcg	1327
Db	119972	cccgatcatctgaagtcctcgaagtcactaagcccaactcagcggcgaaaccgcgaattctcg	120031
Qy	1328	ctgcagagaaatggtttcagaagcatgaatttga	1362
Db	120032	ctgcagagaaatggtttcagaagcatgaatttga	120066

RESULT	4	-				
LOCUS	AX136281/c					
DEFINITION	AX136281	1890 bp	DNA			
ACCESSION	Sequence 203	from Patent	EP1067182.			
VERSION	AX136281					
KEYWORDS	AX136281.1	GI:14272687				
SOURCE	human.					
ORGANISM	Homo sapiens					

REFERENCE
AUTHORS
Ota, T., Isogai, T., Nishikawa, T., Kawai, Y., Sugiyama, T. and
1 (bases 1 to 1890)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Euarctota; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Euarctota; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 203 10-JAN-2001;
Helix Research Institute (JP)
FEATURES Location/Qualifiers
source 1..1890

CDS

```

/note="unamed protein product"
/codon_start=1
/protein_id="CAC39782.1"
/db_xref="GI:14272688"
/translation="MVALGIAATFCGLPILPFAIQIOCYCEEFLNNDCSSEFITY
NCYVAVDMDCKEFMESAGIMYRKSCASSACILASAOYSCFSGKLINVCISCCN
TPLNGSPRPKKRSSASALRPGRLTIIILPLKLASSNHAELKEMPPALLPQSPPT
PHLPE"

```

Query Match	36.5%;	Score 497.6;	DB 6;	Length 1890;
Best Local Similarity	99.08;	Pred. No. 1.5e-89;		
Matches 511; Conservative	0;	Mismatches 4;	Indels 1;	Gaps 1;

QY 848 tcttcctgaggtgatgttgtgcacatggccgcatgtcgtatgcccaaccagattcgga 907
||| || |
Db 1651 TCTGCCAGGGCGTATGTTGTACATTGGCCGTAATGCTGCATGCCCAACCAGATTTCGGA 1592

QY 908 ggaatcatgctgcgcaaacccaagcacgactgtgacgagttcctacttcgcggcgctaca 967
|||||
Db 1591 GGATCATGGCTGCGGCCAACCAGCAGCACTGGACGAGGTCCTACTTCGGGGCGTACA 1532

968 tgatcctcccccctctcgagacgtttttctacccagctcggtcattcaaacccgtcc 1027

Db	1351	TGATGCTGCTCCCTCTCTGGAGAACGTTTTTCTACTCAGCTGCGGTGATCAACCCGGTCC	14172
Qy	1028	tgtatcaagtgctcctcgcaagaatttcgcgcgagtgctgtcagagtgctgtgcgcgc	1087
Db	1471	TGTACACGGGTGCTCTCGACACAGTTTCTGGGGGGGTGTTCTGTGACGGTGTGTTGCTCCGCC	14122
Qy	1088	tgtatcgtagaagacgcgaaccaagaagcgcctgcgcgtacatgcgaattccaccacgcg	11477
Db	1411	TGTGGCTGCGACAGCCACCAACCAAGAAAGGGCTGGCGCTCATATCGCACTCCACACCG	13522
Qy	1148	acagcgcgcgcctcttgtagcaagccgcgctgtgcctcttcgcgtccgcgcgcagtcctcgtcaa	12077
Db	1351	ACACGGCGCGCTTTGTGACACGGCGGTGCTCTTCCGCTCCGGGCCAGTCTCTTGAA	12922
Qy	1208	-ggaagaactgagaagaatttcttaagaacttctcagaagcgagccgaagcccaagtctaag	12667
Db	1291	GGGGAGAACTGAGAAAGATTCTTTAAACACTTTTTCAGAGCGAAGCGCACCCACCACTGTAA	12322
Qy	1267	tcccaagtcattgagtcctcgaggttactaagagcccaactcagcgcgagaaaccagccaattct	13267
Db	1231	TCCCAAGTCATTGAGTGTCTCGAGTCTACTAAGAGCCCAACTCGAGGGCGCAAAACCA6CAATTC	11722
Qy	1327	gctcagaagaatgcttctcagaagcactgaatttga	1362
Db	1171	GCTGCAAGAGATGGTCTTTCAGAGCACTGAAGTTTGA	1136

RESULT	5
LOCUS	ABO41649/c
DEFINITION	ABO41649 1797 bp mRNA ROD
ACCESSION	MUS musculus brain cDNA, clone MNCD-0671.
VERSION	ABO41649
KEYWORDS	ABO41649.1 GI:7670499
SOURCE	fis (full insert sequence).
	Mus musculus (strain:C57Bl), adult female cDNA to mRNA,
	clone:U185c14b, acc:U185c14b, nucdb:0671

ORGANISM

REFERENCE
AUTHORS
1 (sites)
Osada, N., Kusuda, J., Tanuma, R., Ito, A., Hirata, M., Sugano, S. and Hashimoto, K.

TITLE	Isolation of full-length cDNA clones from mouse brain cDNA library
JOURNAL	made by oligo-capping method
REFERENCE	Unpublished (2000)
AUTHORS	2 (Phases 1 to 1997)
TITLE	Hashimoto, K., Osada, N., Kusuda, J. and Sugano, S.
JOURNAL	Direct Submission
	Submitted (13-APR-2000) to the DDBJ/EMBL/Genbank databases.

Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail: khashi@nh.go.jp, URL: <http://www.nih.go.jp/yoken/genbank/>, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181) URL: <http://www.nih.go.jp/yoken/genbank/>

Fish Name: *Surmug mousu*, brain, mch

Vector: pME18S-Fl3

1st strand cDNA was primed with an oligo(dT) primer
[ATGAGCCCTCTTTTCTTTTTTTTTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-Fl3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Library
was constructed by Sugano et al. (University of Tokyo, Institute of
Medical Science). Custom primer used for sequencing (5' end primer
[CTTCTCCTCTAAAGAAGTCGC], 3' end primer
[CGACGTGCACTCGAGCAC]).

A part of this sequence is reported in AU035640.

FEATURES

source	location/qualifiers
	1..1797
	/organism="Mus musculus"

```

/strain="C57BL"
/db_xref="taxon:10090"
/clone="MNCB-0671"
/clone_lib="Sugano mouse brain mncb"
/dev_stage="adult"
/sex="female"
325..750
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA95101.1"
/db_xref="GI:7670500"
/translation="MWVLGIAATPCGLFWLPGTALQIQCYQCEEFQLNDDCSPEETV
NCTVAVDDMCQKEVTEDSAGIMTKKSSASACIIASAGYQSFCSPEKIMSVYCSGNN
TPTLQNGPRPKRKSSASAIRGLTTLTLLFHLALCLAHC"
BASE COUNT      391 a      487 c      511 g      408 t
ORIGIN

```

Query Match	Local Similarity	Conservative	20.9%	Score 285;	DB 10;	Length 1797;
Shes 386;			75.0%;	Pred. No. 4,2e-47;	Mismatches 120;	Indels 9; Gaps
QY 857	gactgattgttgatgcattgacgcatgctgatatgcccacacagatctcgagatcatcg					916
DB 1716	GACTGATTGGTGGAGCGTTGGCCGTGTGTGGATGCCCAATCAATCTCGAGGATCATGG					1657
QY 917	ctgagcgccaaaccccaagcagcgtggaagaggtcctaatttcggcggtacatagtactcc					976
DB 1656	CTGCACACAAACCCAAACATGACTGGACACGACGACTTCTAAGGGCATACTGATCTCC					1597
QY 977	tccctcttcgagagcgttttctcaactcaagtcggtatcaaacccgtctctgtacacg					1036
DB 1596	TGCCCTTCTCTGATACCTCTTCTACCTACGCTGTGATGTCGAACCGTCTCCTTACAACG					1537
QY 1037	tgtctctgcagcagtttcgagcgagtggttcgtgcaggtgctgtgctgcgcgtctgcgtgc					1096
DB 1536	AGTCCTCTCAGCAGTTCCGGAGAGTGTTCGGCAGAGTGCTCTGCTGCGCCTGACTCTGC					1477
QY 1097	agcagcccaacacagagaagcgctgcgtgatacagcgacatccacacagacagcgccc					1156
DB 1476	AGCATGCGCAACCAAGAGAAACGGCACCGGCCGCTTCATCTCCACCAAGACACACCA					1417
QY 1157	gctttgagcagcgccggtgctgctgcgtcc-----cgagcgacgtcctctcaagga					1210
DB 1416	GCTTAGCGCCGACGCCCTCTATTTCTTCTTAGCCTCCCGGCGCAGTAACTTCTCTCAAGA					1357
QY 1211	gaactgagaagatcttcttaagcatttcacagtcagagcgagcc--ccagtltaagt					1267
DB 1356	GAACTTAACAAGGTTTCTTAAAGCATTTCACAGCTGAGGCCAAGCGCTGGACAGGCTAAGC					1297
QY 1268	cccaagtcattgagctcgcagtcacatagagcccaactcagagcgcggaacacagcaatttg					1327
DB 1296	CCGACGCTTGAAGTCCTCGAGCTCACACAGACTGGCTCAAGAACCAACCGCTGGGTCCA					1237
QY 1328	ctgcagaggaatggttttcagagacatgaatttga					1362
DB 1236	CCCCAGAAATAGTTTACAGAGACGAGAAATATGA					1202

TITLE	Secretory protein of membrane protein				
JOURNAL	Patent: EP 1067182-A 620 10-JAN-2001;				
FEATURES	Helix Research Institute (JP)				
source	location/Qualifiers				
	1..587				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	117 a	162 c	140 g	162 t	6 others
ORIGIN					

	Query Match	Similarity	17.4%	Score	236.6	DB	6	Length	587
	Best Local	Similarity	85.8%	Pred.	No.	2.4e-37			
	Matches	296	Conservative	0	Mismatches	43	Indels	6	Gaps
QY	848	tcttcctgagcgtgattgtgtgacattgycgcgtatgctgtgatagccccacgattccga	907						
DB	240	TCCTCCCAAGGCGTGATGTGTGTGACATTTGGCCGTACTGTGATGCCCAACACAGATTGGG	299						
QY	908	ggatcatgctgcgcgcgaaccgaacagatgtgacgaagttccctacttccgcgcgtaca	967						
DB	300	GGATCAACAGCGTGGGGCCAAACCCAAACAGACTGGACAGAGGCTCTACTTCCGGGGCTACA	359						
QY	968	tgtatcccccccctctctcgagacgtttcttaactagctcgtgcatcaaccgcctcc	1027						
DB	360	TGATCTCTCTCCCTTCTGTGGACAGCTTTTCTACCTCAGCTCGGATCATCAACCCGCTTC	419						
QY	1028	tgtaacagtgatcctcgacagcagtttcgcgcgcgtgttctgtaagttgctgtgcgcgcgc	1087						
DB	420	TGTACACGGTCTCTCCCAACAACTTTCGGCGGGTGGTCTGTGAACTTC--TGGCTGCGCC	477						
QY	1088	tgtcgtcgcagcagcgaaccacagagaagcgcctcgcgtctacatgycactccacacag	1147						
DB	478	TGTGCTTTGAACACAGCCAACACGAGAAAGCGCTGAGCC--ACATCGGACATCACAAAC--	534						
QY	1148	acagcgcgcgttgttcgcagcgcgcgtttgctcctcgcgtccgcgc	1192						
DB	535	-GAAAGCGCCCTTTGGGCAACGCCCGGTGGTTTTGTGATCCGGGCG	578						

RESULT	7	
XX	E11480	
XX	E11480	standard; RNA; HUM; 1254 BP.
XX	E11480;	
XX	E11480.1	
XX	08-OCT-1997 (Rel. 52, Created)	
DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)	
XX	CDNA encoding human neurotensin receptor protein.	
DE		
XX	JP 1996143597-A/1.	
XX		
OS	Homo sapiens (human)	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia	
OC	Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
XX		
RN	[1]	
RP	1-1254	
RA	Onda H., Oogi K., Hinuma K., Masuo Y.;	
RT	"HUMAN NEUROTENSIN RECEPTOR PROTEIN, ITS PRODUCTION AND USE";	
RL	Patent number JP1996143597-A/1, 04-JUN-1996.	
RL	TAKEIDA CHEM IND LTD.	
XX		
OS	Homo sapiens (human)	
CC	PN JP 1996143597-A/1	
CC	PD 04-JUN-1996	
CC	PF 04-NOV-1994 JP 1994289882	
CC	PI ONDA HARUO, OOGI KAZUHIRO, HINUMA KUNIJU, MASUO YOSHINORI	
CC	PC OKJ14/72, A61K38/00, A61K45/00, A61K45/00, A61K45/00,	
CC	PC A61K45/00,	

RESULT	6			
AXI36698				
LOCUS	AXI36698	587 bp	DNA	PAT 30-MAY-2001
DEFINITION	Sequence 620 from Patent EP1067182.			
ACCESSION	AXI36698			
VERSION	AXI36698.1	GI:14273102		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 587)			
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.			

xx	
RP	1-1254
RA	Onda H., Oogi K., Hinuma K., Masuo Y.;
RT	"HUMAN NEUTROTESIN RECEPTOR PROTEIN, ITS PRODUCTION AND USE";
RL	Patent number JP1996143597-A/1, 04-JUN-1996.
XL	TAKEDA CHEM IND LTD.
xx	
CC	OS Homo sapiens (human)
CC	PN JP 1996143597-A/1
CC	PD 04-JUN-1996
CC	PF 24-NOV-1994 JP 1994289882
CC	PI ONDA HARUO, OGGI KAZUHIRO, HINUMA KUNITJI, MASUO YOSHINORI
CC	C07K14/772,A61k38/00,A61k45/00,A61k45/00,A61k45/00,A61k45/00,
CC	PC A61k45/00,

VERSION AL357872.15 GI:11414584
 SOURCE HTGS_HTSASL1.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 367980)
 TITLE Plumb, B.
 JOURNAL Direct Submission
 Submitted (08-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Nov 28, 2000 this sequence version replaced gi:114340280.
 COMMENT
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: BA308BP
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: piasmld; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 346090 bases at least Q40
 Consensus quality: 353078 bases at least Q30
 Consensus quality: 356859 bases at least Q20
 Insert size: 362580; sum-of-contigs
 Insert size: 177412; 7.0% error; agarose-fp
 Quality coverage: 4.17x in Q20 bases; sum-of-contigs Quality
 coverage: 9.44x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 55 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 3371: contig of 3371 bp in length
 * 3372 3471: gap of 100 bp
 * 3472 6278: contig of 2807 bp in length
 * 6279 6378: gap of 100 bp
 * 6379 11841: contig of 5463 bp in length
 * 11842 11941: gap of 100 bp
 * 11942 17529: contig of 5588 bp in length
 * 17530 17629: gap of 100 bp
 * 17630 25254: contig of 7625 bp in length
 * 25255 25354: gap of 100 bp
 * 25355 30177: contig of 4823 bp in length
 * 30178 30277: gap of 100 bp
 * 30278 33210: contig of 2933 bp in length
 * 33211 33310: gap of 100 bp
 * 33311 43615: contig of 10305 bp in length
 * 43616 43715: gap of 100 bp
 * 43716 47706: contig of 3991 bp in length
 * 47707 47806: gap of 100 bp
 * 47807 51335: contig of 3529 bp in length
 * 51336 51435: gap of 100 bp
 * 51436 53590: contig of 2155 bp in length
 * 53591 53690: gap of 100 bp
 * 53691 56178: contig of 2468 bp in length
 * 56179 56278: gap of 100 bp
 * 56279 59223: contig of 2945 bp in length
 * 59224 59323: gap of 100 bp
 * 59324 63598: contig of 4275 bp in length
 * 63599 63698: gap of 100 bp
 * 63699 77580: contig of 13882 bp in length
 * 77581 77680: gap of 100 bp
 * 77681 81030: contig of 3350 bp in length
 * 81031 81130: gap of 100 bp
 * 81131 90501: contig of 9371 bp in length

* 90502 90601: gap of 100 bp
 * 90602 98709: contig of 8108 bp in length
 * 98710 98809: gap of 100 bp
 * 98810 107616: contig of 8807 bp in length
 * 107617 107716: gap of 100 bp
 * 107717 114867: contig of 7151 bp in length
 * 114868 114967: gap of 100 bp
 * 114968 123825: contig of 8858 bp in length
 * 123826 123925: gap of 100 bp
 * 123926 126983: contig of 3058 bp in length
 * 126984 127083: gap of 100 bp
 * 127084 130970: contig of 3887 bp in length
 * 130971 131070: gap of 100 bp
 * 131071 136942: contig of 5872 bp in length
 * 136943 137042: gap of 100 bp
 * 137043 144404: contig of 7362 bp in length
 * 144405 144504: gap of 100 bp
 * 144505 158206: contig of 13702 bp in length
 * 158207 158306: gap of 100 bp
 * 158307 164782: contig of 6476 bp in length
 * 164783 164882: gap of 100 bp
 * 164883 169747: contig of 4865 bp in length
 * 169748 169847: gap of 100 bp
 * 169848 174377: contig of 4530 bp in length
 * 174378 174477: gap of 100 bp
 * 174478 177983: contig of 3506 bp in length
 * 177984 178083: gap of 100 bp
 * 178084 181525: contig of 3442 bp in length
 * 181526 181625: gap of 100 bp
 * 181626 190997: contig of 9372 bp in length
 * 190998 191097: gap of 100 bp
 * 191098 195091: contig of 3994 bp in length
 * 195092 195191: gap of 100 bp
 * 195192 198115: contig of 2924 bp in length
 * 198116 198215: gap of 100 bp
 * 198216 206597: contig of 8382 bp in length
 * 206598 206697: gap of 100 bp
 * 206698 220397: contig of 13700 bp in length
 * 220398 220497: gap of 100 bp
 * 220498 222979: contig of 2482 bp in length
 * 222980 223079: gap of 100 bp
 * 223080 225370: contig of 2291 bp in length
 * 225371 225470: gap of 100 bp
 * 225471 229357: contig of 3887 bp in length
 * 229358 229457: gap of 100 bp
 * 229458 231791: contig of 2334 bp in length
 * 231792 231891: gap of 100 bp
 * 231892 235072: contig of 3181 bp in length
 * 235073 235172: gap of 100 bp
 * 235173 241381: contig of 6309 bp in length
 * 241382 241481: gap of 100 bp
 * 241482 258412: contig of 16931 bp in length
 * 258413 258512: gap of 100 bp
 * 258513 267906: contig of 9394 bp in length
 * 267907 268006: gap of 100 bp
 * 268007 280454: contig of 12448 bp in length
 * 280455 280554: gap of 100 bp
 * 280555 292423: contig of 11869 bp in length
 * 292424 292523: gap of 100 bp
 * 292524 298279: contig of 5756 bp in length
 * 298280 298379: gap of 100 bp
 * 298380 301682: contig of 3303 bp in length
 * 301683 301782: gap of 100 bp
 * 301783 305013: contig of 3331 bp in length
 * 305014 305113: gap of 100 bp
 * 305114 308387: contig of 3374 bp in length
 * 308388 308487: gap of 100 bp
 * 308488 311059: contig of 2572 bp in length
 * 311060 311159: gap of 100 bp
 * 311160 313655: contig of 2496 bp in length
 * 313656 313755: gap of 100 bp
 * 313756 316142: contig of 2387 bp in length
 * 316143 316242: gap of 100 bp


```

repeat_region      916. .995
                    /note="2 copies 40 mer 93% conserved"
repeat_region      1033. 1112
                    /note="4 copies 20 mer 81% conserved"
repeat_region      3871. 3906
                    /note="18 copies 2 mer aa 88% conserved"
misc_feature       4585. 5025
                    /note="match: GSS: Em: A2537553"
repeat_region      4699. 4766
                    /note="2 copies 34 mer 98% conserved"
repeat_region      5375. 5553
                    /note="MIR repeat: matches 43. 236 of consensus"
misc_feature       5645. 7284
                    /note="Cpg island"
                    /evidence=not_experimental
mRNA               join(6239. 6272, 20078. 20969, 22396. 22486, 24131. 24252,
                    24783. 24894, 28653. 28807, 30099. 30294, 31464. 32260,
                    32649. 32797, 35469. 36012)
                    /gene="SLC21A12"
                    /note="match: CDNAS: Em: AF104334
                    match: ESTs: Em: A1022939 Em: A1347130 Em: A574753
                    Em: A280994 Em: A1348176 Em: A534529 Em: A1803717"
                    /evidence=not_experimental
                    /product="ba93b14.1.1 (solute carrier family 21 (organic
                    anion transporter) member 12, isoform 1)"
gene               6239. 36012
                    /gene="SLC21A12"
                    /note="MIR repeat: matches 8. 246 of consensus"
repeat_region      9156. 9197
                    /note="7 copies 6 mer aggtgt 90% conserved"
repeat_region      9204. 9251
                    /note="8 copies 6 mer caccct 89% conserved"
repeat_region      12696. 13185
                    /note="MLT2CB repeat: matches 1. 501 of consensus"
repeat_region      13745. 13834
                    /note="2 copies 45 mer 96% conserved"
repeat_region      15721. 15816
                    /note="3 copies 32 mer 77% conserved"
repeat_region      15722. 15799
                    /note="6 copies 13 mer 75% conserved"
repeat_region      15726. 15799
                    /note="37 copies 2 mer cc 75% conserved"
repeat_region      15932. 15969
                    /note="19 copies 2 mer tg 81% conserved"
repeat_region      15934. 15969
                    /note="9 copies 4 mer tgtg 83% conserved"
repeat_region      16659. 16762
                    /note="MIR repeat: matches 29. 138 of consensus"
repeat_region      16743. 16800
                    /note="L2 repeat: matches 2625. 2687 of consensus"
repeat_region      17515. 17762
                    /note="LTPA5 repeat: matches 5898. 6145 of consensus"
repeat_region      17936. 18075
                    /note="MLT1C repeat: matches 53. 194 of consensus"
repeat_region      18516. 18653
                    /note="MLT1B repeat: matches 303. 440 of consensus"
repeat_region      18682. 18860
                    /note="MLT1B repeat: matches 1. 181 of consensus"
repeat_region      18903. 19126
                    /note="4 copies 56 mer 96% conserved"
repeat_region      join(20078. 20969, 22396. 22486, 24131. 24252, 24783. 24894,
                    28653. 28807, 30099. 30294, 31464. 31629, 31731. 31903,
                    32196. 32260, 32649. 32797, 35469. 36012)
                    /gene="SLC21A12"
                    /note="match: CDNAS: Em: AB031051
                    match: ESTs: Em: A1022939 Em: A280994 Em: A574753
                    Em: A1347130 Em: A5027513 Em: A1803717 Em: A287831"
                    /evidence=not_experimental
                    /product="ba93b14.1.2 (solute carrier family 21 (organic
                    anion transporter) member 12, isoform 2)"
mRNA               join(20174. 20969, 22396. 22486, 24131. 24252, 24783. 24894,
                    28653. 28807, 30099. 30294, 31464. 31629, 31731. 31903,
                    32196. 32260, 32649. 32797, 35469. 36012)

```

CDS

```

/gene="SLC21A12"
/note="there is a 1 base insertion at position 31786 as
compared with the Em: AF104334 and therefore results in a
difference between protein sequences.
match: proteins: Tr:062523 Tr:090135"
/codon_start=1
/evidence=not_experimental
/product="ba93b14.1.1 (solute carrier family 21 (organic
anion transporter) member 12, isoform 1)"
/db_xref="GI:11121398"
/translation="MPHLQGDKPLRPPSPNSAMENGLDHTPPSRASPTPLPSGL
RSAHSPIDTSKQPLCOLMAEKHGARTHEVRVSAQSVACQWMAFAPCLQVLTNP
KGLIFLCAAFLEGMTVNGFINVTISLERVYLSHYSQGLASADIAACLTFFV
SYFGSGHKRPFWLMGVLMGTGSLVALPHEFTGTYREVELDGVTRCPANPGAVCAD
STGSLSRLOLVFPMGLHGVGATPLTICVTIDENKVSQSPVYIAITFYTAALICP
AAGLIGGALLNITYEMGRRELTPRESPIVGMWVFSLGCAAFPTAVPILIGPRQ
LPGSORVAVRAAMHOLKSSRGEASNPFGKTRDLPUSLWLLKNPFIILCLAG
ATEPTLITGMSSTSPKFLSOFSLASAEATLFGYLVNPPAGGGTFLGFFVKRLR
GSAVIRKCLCTVAVSLGILVFSLHCPVSMAGVTASVGGSLPEGHNLITAPCNAC
SCOPEHNSPVGSDGLMYFSLCHAGCPAATEVNDGKVSAAARPCPLDYGKGP
CLPLVIGAIVGLPCTETVAVSLRIFPLVAMPLOGNALQLVRESPEFMSYSL"
join(20174. 20969, 22396. 22486, 24131. 24252, 24783. 24894,
28653. 28807, 30099. 30294, 31464. 31629, 31731. 31903,
32196. 32260, 32649. 32797, 35469. 36012)
/gene="SLC21A12"
/note="match: proteins: Tr:090135"
/codon_start=1
/evidence=not_experimental
/product="ba93b14.1.2 (solute carrier family 21 (organic
anion transporter) member 12, isoform 2)"
/protein_id="CAC14920.1"
/db_xref="GI:11121399"
/translation="MPHLQGDKPLRPPSPNSAMENGLDHTPPSRASPTPLPSGL
RSAHSPIDTSKQPLCOLMAEKHGARTHEVRVSAQSVACQWMAFAPCLQVLTNP
KGLIFLCAAFLEGMTVNGFINVTISLERVYLSHYSQGLASADIAACLTFFV
SYFGSGHKRPFWLMGVLMGTGSLVALPHEFTGTYREVELDGVTRCPANPGAVCAD
STGSLSRLOLVFPMGLHGVGATPLTIGVLTIDENKVSQSPVYIAITFYTAALICP
AAGLIGGALLNITYEMGRRELTPRESPIVGMWVFSLGCAAFPTAVPILIGPRQ
LPGSORVAVRAAMHOLKSSRGEASNPFGKTRDLPUSLWLLKNPFIILCLAG
ATEPTLITGMSSTSPKFLSOFSLASAEATLFGYLVNPPAGGGTFLGFFVKRLR
GSAVIRKCLCTVAVSLGILVFSLHCPVSMAGVTASVGGSLPEGHNLITAPCNAC
SCOPEHNSPVGSDGLMYFSLCHAGCPAATEVNDGKVSAAARPCPLDYGKGP
CLPLVIGAIVGLPCTETVAVSLRIFPLVAMPLOGNALQLVRESPEFMSYSL"
join(20174. 20969, 22396. 22486, 24131. 24252, 24783. 24894,
28653. 28807, 30099. 30294, 31464. 31629, 31731. 31903,
32196. 32260, 32649. 32797, 35469. 36012)
/gene="SLC21A12"
/note="match: proteins: Tr:090135"
/codon_start=1
/evidence=not_experimental
/product="ba93b14.1.2 (solute carrier family 21 (organic
anion transporter) member 12, isoform 2)"
/protein_id="CAC14920.1"
/db_xref="GI:11121399"
/translation="MPHLQGDKPLRPPSPNSAMENGLDHTPPSRASPTPLPSGL
RSAHSPIDTSKQPLCOLMAEKHGARTHEVRVSAQSVACQWMAFAPCLQVLTNP
KGLIFLCAAFLEGMTVNGFINVTISLERVYLSHYSQGLASADIAACLTFFV
SYFGSGHKRPFWLMGVLMGTGSLVALPHEFTGTYREVELDGVTRCPANPGAVCAD
STGSLSRLOLVFPMGLHGVGATPLTIGVLTIDENKVSQSPVYIAITFYTAALICP
AAGLIGGALLNITYEMGRRELTPRESPIVGMWVFSLGCAAFPTAVPILIGPRQ
LPGSORVAVRAAMHOLKSSRGEASNPFGKTRDLPUSLWLLKNPFIILCLAG
ATEPTLITGMSSTSPKFLSOFSLASAEATLFGYLVNPPAGGGTFLGFFVKRLR
GSAVIRKCLCTVAVSLGILVFSLHCPVSMAGVTASVGGSLPEGHNLITAPCNAC
SCOPEHNSPVGSDGLMYFSLCHAGCPAATEVNDGKVSAAARPCPLDYGKGP
CLPLVIGAIVGLPCTETVAVSLRIFPLVAMPLOGNALQLVRESPEFMSYSL"
join(20174. 20969, 22396. 22486, 24131. 24252, 24783. 24894,
28653. 28807, 30099. 30294, 31464. 31629, 31731. 31903,
32196. 32260, 32649. 32797, 35469. 36012)

```

Query Match

Best Local Similarity 57.1% Pred. No. 1.4e-20; Length 114642;

Matches 274; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

```

QY 106 ggtgactgataccttgatggccttcgtgggaacagccaccattcgggtcacc 165
DB 73131 ggtgactgataccttgatggccttcgtgggaacagccaccattcgggtcacc 165

```


181 aaagataacttgcagaagaggtgacagaccacatgltgatttgccttgctcggacatc 240

```

source
1..170695
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-496H23"
/clone_1b="RPCT-11.2"
1..50299
misc-feature
1..note="assembly-fragment:02503"
50400..170695
note="assembly-fragment:04092"
04092..
misc-feature
1..

```

